

Linkage Technologies Which Enhance the Utility of Task-Based Occupational Analysis

William J. Phalen, Jimmy L. Mitchell
Institute for Job & Occupational Analysis
San Antonio, Texas, U.S.A.

Winston Bennett, Jr.
Technical Training Research Division
Armstrong Laboratory, Human Resources Directorate
Brooks Air Force Base, Texas, U.S.A.

Darryl K. Hand
Metrica, Inc.
San Antonio, Texas
U.S.A.

19990608 120

Occupational researchers and analysts continue to search for new ways to meet the needs of the manpower, personnel, and training (MPT) communities they serve. Recent efforts seem to be moving away from task-based occupational analysis toward more global measures of knowledges, skills, and abilities (KSAs) and job content. It is alleged that traditional task-based occupational analysis is too labor intensive, too costly, too cumbersome, and too static to meet the emerging and rapidly changing needs of a business process- and team-based approach to the organization of work activities (Casio, 1995; May, 1996 a-c). More and more, the trend has been toward relying on the judgments and opinions of a few subject matter experts (SMEs) on a handful of broad work dimensions or requirements to make important policy decisions.

All the while, there lies largely untapped a treasure trove of data and analysis techniques in the "old" occupational measurement system (loosely identified as "CODAP") that would address many knotty MPT problems with greater precision and in greater depth than the newer approaches, if only a little mental effort were applied. The great advantage of gathering task-level data is that it can be aggregated and linked to other data bases, such as personnel files, in innumerable ways to address specific problems that were not even envisioned when the data were gathered.

This paper will address four areas where existing and proposed linkage technologies can be applied to existing or updated occupational and personnel data bases to address largely unmet needs of the MPT communities. The four areas are empirically derived task difficulty measures within and across career fields, linking of specific learning abilities to job performance, weapon system-to-task linkage technology, and knowledge-to-task linkage technology.

Empirical Derivation of Benchmarked Task Difficulty Measures

The existing procedure for estimating task difficulty is highly subjective and rather unreliable, in that a group of randomly selected SMEs cannot be unbiased and skilled raters for all the tasks in a USAF Job Inventory and do not clearly perceive an "intrinsic" level of difficulty in a task such as would lead to high interrater agreement. The low average interrater agreement statistic (R_{11}) routinely obtained (.10 - .30 in most cases) indicates generally poor agreement. The very high projected interrater agreement (R_{KK}) statistics that are obtained are primarily a function of the large number of raters used. The process of benchmarking tasks across Air Force Specialties (AFSs) is even more subjective and more suspect, in that raters must rate and rank order tasks from AFSs other than their own and use a benchmarking scale that has been shown to be biased by the selection of tasks used to anchor the various benchmark scale points. Furthermore, since the benchmarking is accomplished separately for mechanical, general/administrative, and electronic aptitude (MAGE) areas, a further benchmarking has to be accomplished to create a single benchmarked difficulty. This has been done mathematically, based on a number of assumptions.

Because the entire benchmarking process is so complex and labor-intensive, it has not been feasible to repeat it, even when a considerable number of AFSs have changed substantially. Thus, recent attempts to update the benchmarking have resorted to updating current task difficulty indices, based on the assumption that changes in task difficulty values in an AFS can be transformed directly into changes in benchmarked difficulty values without reference to all the changes going on simultaneously in all other AFSs. These estimates are thrice-removed from the original task difficulty which, as has been pointed out, already has its own problems.

A better solution would seem to be one which uses a number of readily available measures that are empirical, rather than subjective, absolute and universal, rather than relative and AFS-specific, and which clearly represent the way tasks are perceived, in terms of task difficulty, in the real world of work.

Thus, within AFSs, it would seem that the difficulty of a task could be viewed primarily as a function of four "experience" variables: average time in service (TAFMS) of people performing the task; the average time in the career field (TICF); the average skill level (based on duty AFS code); and the average paygrade of those who perform the task (as computed by the CODAP GRPAVG program). These "experience" variables indirectly include training, in that on-the-job training (OJT) time adds to the average values, and technical school training time also adds months to the average TAFMS and TICF variables. To some extent, aptitude may enter in here, too, but aptitude is primarily a variable which distinguishes between AFSs and does so in terms of the four MAGE variables tested in the Armed Services Vocational Aptitude Battery (ASVAB). Differences in task difficulty between AFSs would be determined by the average M, A, G, & E values of those who perform each task. All four aptitude variables would be used for each AFS because: (1) comparability is desired across all AFSs (hence every task should have the same set of aptitude scores), and (2) it should not be assumed that every task in an electronic AFS, for example, has an electronic aptitude requirement (some tasks may require more mechanical, general, or administrative aptitude). It should be noted at this point that the four experience and the four MAGE variables have the same meaning across all specialties (e.g., a "TAFMS" of nine months means the same for all AFSs; so, too, would an aptitude of "E-50"). Standardization of each variable across all AFSs would allow for a synthetically weighted benchmarked task difficulty composite score to be generated for each task, using equal or unequal weighting of the eight variables. On the other hand, the eight variables could be used as predictors of the existing benchmarked task difficulty. If the R^2 is high, it should replace the existing system. If R^2 is not high, validation of the proposed system could be accomplished by selecting pairs of tasks on which the two systems yielded widely divergent values and have knowledgeable SMEs rank order the pairs of different tasks. A simple χ^2 test would determine which system was more accurate.

In the benchmarking system proposed here, the entire benchmarking process could be redone immediately whenever an AFS is resurveyed. The benchmarked occupational learning difficulty (BOLD) index could also be computed for each AFS using an equation that proved to be very promising as a measure of job difficulty in the Computer-Administered Survey Software (CASS) project conducted by the Armstrong Laboratory (Albert, et al., 1994).

Linking of Specific Learning Abilities to Job Performance

Within the last several years, airmen entering technical school training in 15 AFSs have been administered a learning abilities test covering six cognitive processes applied to verbal, quantitative, and spatial context. Dr. Linda Sawin and her associates at the Armstrong Laboratory are now in the process of validating these measures against technical school grades. There is great potential for following these airmen into the field and assessing their job performance after six to 12 months on the job, and again after two to three years. First of all, it could be determined

whether there are differences in the tasks being performed by airmen who are high or low in various learning abilities and to what extent more general aptitudes from the ASVAB (MAGE) and school grades are accounting for the same variance. An airman's job proficiency could be assessed with a survey instrument administered to the airman's supervisor and peers, using a procedure similar to one successfully used in the automated test outline (ATO) project to assess various aspects of job performance and job knowledge.

The learning abilities, MAGE scores, and school grades could be used to predict overall or various aspects of job performance and job knowledge for all airmen in each AFS, and for groups of airmen identified as belonging to specific job types (as identified by AFOMS occupational analysts). This would not only provide a job-based validation of the learning abilities, but also assess their individual contributions to success in specific jobs in each AFS over and above other aptitude measures and school grades. If their contribution is significant, they will fill the "ability" gap that has not been adequately filled by any existing "ability" paradigm.

Weapon System-to-Task Linkage

The background section of the standard Air Force Job Inventory is a rich source of informative variables whose value is greatly enhanced when linked with the respondent's task data and aggregated across like respondents. One set of variables that is included in many equipment-oriented AFSs is a list of weapon systems. Respondents are asked to check all weapon systems on which they work. They could be asked, also, to check all weapon

systems with which they are familiar. Better yet, they might be asked to rate their level of familiarity or involvement with these systems. This information would provide the prerequisite data for a powerful new technology that is able to not only link tasks to specific weapons systems, but also to estimate the similarity of weapon systems relative to the performance of individual tasks. The almost unthinkable number of linkages that are required to produce all possible similarities for a large number of weapons systems and an even larger number of tasks can be accomplished efficiently by a shorthand coding system that allows raters to indicate weapon system linkage to a task and inter-weapon system similarity regarding task performance with letter codes and hyphens, such that, for example, a coded listing of six weapon systems for a task would yield $7(6)/2 = 21$ pairwise similarity relationships among the weapon systems listed. All of these relationships would be generated by the computer from the rater's relatively simple exercise. The computer would then use these similarity relationships to generate an exponentially large number of additional, indirect similarity relationships among weapons for individual tasks by means of a networking algorithm that uses transitivity to identify the additional relationships, many of which would involve comparisons that were never directly made by the raters. For example, even if weapon systems "A" and "B" were judged to be similar, as were weapon systems "B" and "C" (in the performance of a specific task), it could be inferred that weapon systems "A" and "C" are similar.

The resultant matrix of weapon system-to-task similarity indices would have multiple uses. They could be used to assure the assignment of qualified personnel from one weapon system to another or needed cross-training, if such an assignment is made. They could be used to identify the most qualified personnel for assignment to new weapon systems that have task-level commonalties with one or more existing systems. They could, of course, be used to determine personnel and training requirements for new weapon systems, even before they are developed, based on their known commonality with existing weapon systems. Perhaps, most importantly, this technology would make possible the expansion of a standard job inventory to a weapon system-based job inventory after a standard job inventory has been administered. This would be accomplished by breaking out tasks that are performed differently on weapon systems "A" and "B" versus "C," "D," and "E," for example, into two subtasks. An individual inventory respondent's time spent values on these subtasks would be allocated according to the time they indicated they spend on each weapon system in the background section of their inventory. The weapon system-to-task linkage process is described in more detail in Phalen & Mitchell (1993), and an application of the process is described in great detail in Bennett & Phalen (1995).

Knowledge-to-Task Linkage Technology

Numerous efforts have been made to link knowledges to tasks, some more successful than others. Even assuming that the lists of knowledges and tasks are well constructed, a major logistical problem always remains. If the knowledges and tasks are appropriately specific, the lists are long and the matching, for example, of 100 knowledges with 1,000 tasks would potentially require 10,000 matchings. If more abbreviated lists are developed, the knowledges and/or tasks will be too general to be as useful as desired. One possible way to circumvent the problem is to have the job incumbent rate the task list in the normal way, but rate each knowledge (as a background item) according to whether he or she uses it on the job. This would reduce the potential number of responses in the previous example from $100 \times 1,000 = 10,000$ to $100 + 1,000 = 1,100$.

The problem now is how to link tasks to knowledges which have been rated only at the job level. Although still in the developmental stage, an algorithm has been developed for doing this which shows some promise of eventually becoming a useful tool. The algorithm differentially weights negative and positive linkage information. The heaviest weight, which is a negative weight, is given to a "1,0" relationship, in which a respondent says he or she performs the task but does not use the knowledge on the job. This is direct "lack of linkage" information. On the other hand "1,1" items are given a positive weight only to the extent that there are "0,0" items to match them. The reason for this is that some tasks and knowledges may be performed/used by almost everyone and no linkage between that knowledge and that task can be inferred from the large number of "1,1" responses. The maximum positive linkage value would occur if half the respondents said they performed the task and used the knowledge, and the other half said they did not perform the task and did not use the knowledge. A "0,1" pairing of responses, as well as the excess of "1,1" and "0,0" responses that are not matches, i.e., $N_T - \min(n_{11}, n_{00})$ have a slight negative impact (like "noise") in the algorithm. The algorithm and an example application can be obtained by contacting the senior author.

The algorithm will be most successful in situations in which most of the knowledges and task have been checked or rated by 25% to 75% of the respondents. Outside these limits, its usefulness decreases rapidly. But even in most cases where a clear linkage cannot be established between a single knowledge and a single task, the number of tasks potentially linked to a knowledge is small enough that it would not be an overwhelming job for SMEs to review the short lists and make the final linkages. The lists can be generated to show which tasks are most closely linked to a specific knowledge or which knowledges are most closely linked to a specific task.

Once linkages have been computed between every knowledge and every task, the profiles of task linkages for each knowledge can be used to cluster knowledges across tasks, resulting in knowledge clusters that are linked to a common set of tasks (that can also be displayed). Contrariwise, the matrix of linkage values can be transposed and the profiles of task linkages across knowledges can be used to cluster tasks across knowledges, resulting in task clusters that are linked to a common set of knowledges (that can also be displayed).

Software has already been developed to run the linkage algorithm on occupational survey data and produce the products described above. Although further research and development of the knowledge-to-task linkage technology is required, it is hoped that an operational system will be available by the time of the next IMTA conference. We will soon be testing the technology on several appropriate data sets that were kindly furnished to us by the Canadian Forces Directorate of Personnel Planning.

Conclusions

Four linkage technologies have been outlined in this paper. Each is based on available technology or technology-in-progress. Each has the potential for addressing important MPT needs that are not currently being addressed as well as these technologies would permit. There are actually many more data linkage possibilities that could not be addressed within the limitations of this paper but are readily available to those who have a mind to look for them. For all of these technologies, a whole range of possibilities exists for multi-level analysis, since task-level data, including data linked to tasks, can be readily aggregated to any level of specificity that MPT managers at any level might find informative and useful. The same cannot be said for the global constructs and dimensions being proposed as replacements for detailed task information.

References

- Albert, W.G., Phalen, W.J., Selander, D.M., Dittmar, M.J., Tucker, D.L., Hand, D.K., Weissmuller, J.J. & Rouse, I.F. (1994, October). Large-scale laboratory test of occupational survey software and scaling procedures. In the symposium, Bennett, W. Jr., Chair, Training needs assessment and occupational measurement: Advances from recent research. *Proceedings of the 36th Annual Conference of the International Military Testing Association*. Rotterdam, The Netherlands: European Members of the IMTA.
- Bennett, W., Jr., & Phalen, W.J. (1995). Development and preliminary results from an approach for linking tasks and knowledge items to different weapon systems. In the symposium, R.B. Gould (Chair), Issues and advances in task-based occupational research and development for manpower, personnel, and training. *Proceedings of the 37th annual conference of the International Military Testing Association*, Toronto, Canada.
- Cascio, W.F. (1995). Whither industrial and organizational psychology in a changing world of work? *American Psychologist*, 50, 928-939.
- May, K.E. (1996a). Work in the 21st century: Implications for job analysis. *The Industrial-Organizational Psychologist*, 33(4), 98-100.
- May, K.E. (1996b). Work in the 21st century: Implications for performance management. *The Industrial-Organizational Psychologist*, 34(1), 23-24.
- May, K.E. (1996c). Work in the 21st century: Implications for compensation. *The Industrial-Organizational Psychologist*, 34(2), 73-77.
- Phalen, W.J., & Mitchell, J.L. (1993). New approaches for increasing information value of individual response data. In the symposium, Organizational Analyses and Research: Policy Issues, Modeling, and Future Technologies (W.R. Bennett, Jr. Chair) *Proceedings of the 35th Annual Conference of the Military Testing Association*. Williamsburg, VA: U.S. Coast Guard Headquarters, Occupational Analysis Program.

 [Back to Table of Contents](#)

INTERNET DOCUMENT INFORMATION FORM

A . Report Title: Linkage Technologies Which Enhance the Utility of Task-Based Occupational Analysis

B. DATE Report Downloaded From the Internet: 06/04/99

C. Report's Point of Contact: (Name, Organization, Address, Office Symbol, & Ph #):
Navy Advancement Center
ATTN: Dr. Grover Diel (850) 452-1615
Pensacola, FL

D. Currently Applicable Classification Level: Unclassified

E. Distribution Statement A: Approved for Public Release

F. The foregoing information was compiled and provided by:
DTIC-OCA, Initials: __VM__ Preparation Date 06/04/99

The foregoing information should exactly correspond to the Title, Report Number, and the Date on the accompanying report document. If there are mismatches, or other questions, contact the above OCA Representative for resolution.